Methodology for monitoring progress towards the global nutrition targets for 2025

Technical report

by the

WHO-UNICEF Technical Expert Advisory Group on

Nutrition Monitoring (TEAM)

June 2017

WHO-UNICEF Technical Expert Advisory Group on Nutrition Monitoring (TEAM)

Contents

Introduction

In 2012 the World Health Assembly, in resolution WHA65.6, endorsed a *Comprehensive Implementation Plan on Maternal, Infant and Young Child Nutrition*, which specified six global nutrition targets for 2025:

- ➢ achieve a 40% reduction in the number of children under five who are stunted;[1]
- ➢ achieve a 50% reduction of anaemia in women of reproductive age;
- ➢ achieve a 30% reduction in low birth weight;
- ➢ ensure that there is no increase in childhood overweight;
- ➢ increase exclusive breastfeeding (EBF) rates in the first 6 months up to at least 50%;
- ➢ reduce and maintain childhood wasting to less than 5%.

In 2015, the United Nations General Assembly agreed on the Sustainable Development Goals (SDGs) as part of the post-2015 development agenda. The second of the seventeen proposed SDGs is "End hunger, achieve food security and improve nutrition." This goal promotes sustainable agriculture and includes the target to end all forms of malnutrition by 2030, embracing WHO targets on stunting, wasting and overweight in children under five years of age, and addressing the nutritional needs of adolescent girls, pregnant and lactating women, and older persons.

Awareness of the importance of nutrition is increasing; its key role for sustainable development is acknowledged by governments and stakeholders around the world. To monitor progress on commitments and identify opportunities for action, common rules to classify countries based on their progress towards commitments are needed.

For the six nutrition targets, there has been discussion on how to best capture country progress towards targets, while encouraging further progress and acknowledging efforts.

There is growing agreement that comparable and equitable global monitoring across countries has a distinct purpose to countries monitoring achievement of national targets. Global monitoring sustains dialogue between countries as it provides a common yardstick to help understand two complementary concepts – the magnitude of the problem and progress achieved over time in resolving it.

Global goals and targets do not always easily translate into country-specific ones. Nevertheless, their monitoring can help inform tracking progress toward achieving national targets. Global and country target monitoring share the same challenges regarding the availability of timely, high-quality data, which is a problem in many countries. For example, relatively few high-income countries have national data on exclusive breastfeeding (EBF) or stunting. Even when data seem plentiful, monitoring efforts can be hampered by inconsistent indicator definitions.

---

[1] Note the target for stunting is the only one that is based on numbers affected; target and progress assessment thus take into consideration both the change in prevalence and a country's population growth.

Even though, for most of the target indicators, it is unrealistic to expect changes in outcome and impact within a year or two, let alone data reflecting change, there is a high demand for reporting on progress. Indeed, reporting on progress in achieving targets started soon after their endorsement, well before 2012 baseline data were available in many countries. Using global databases from WHO and UNICEF, country classifications were presented in the Global Nutrition Report (GNR) 2014-2016 publications.[2,3,4] The GNR classification rules have changed over the years.

The present report examines the approach used for monitoring progress towards targets and discusses the technical complexities experienced in assessing progress in individual countries in a consistent and comparable fashion.

WHO and UNICEF collaborate in global monitoring and reporting on nutrition indicators. In 2015, they established a joint WHO-UNICEF Technical Expert Advisory Group on Nutrition Monitoring (TEAM). TEAM advises on improving nutrition monitoring at all levels by sharing experiences and developing harmonized standards, tools and approaches.

In February 2016, the joint TEAM secretariat presented unresolved issues concerning the tracking rules to the TEAM. The group proposed further analyses to support recommended rules, which led to a new background document being presented during the September 2016 meeting. This report summarizes the main issues, the evolution in tracking, and final recommendations. Recommendations will evolve based on experience and common agreement that will ensure comparability over time and the best use of available data.

Development of the monitoring rules to assess country progress towards the global nutrition targets

In May 2014, WHO proposed a simple approach for monitoring four of the six targets (stunting, wasting, overweight and anaemia), based on their global target definitions to present to the World Health Assembly and to be used for the first GNR. In the same year, the GNR included these rules with some modifications and including rules for exclusive breastfeeding.[2]
The rules in the next two GNR editions evolved taking into account understanding about the on- or off-course classification of countries with respect to progress towards the global targets. A new methodology was proposed for exclusive breastfeeding with recommended country targets. Table 1 summarizes the various versions of these rules.

---

[2] International Food Policy Research Institute. Global Nutrition Report 2014: Actions and Accountability to Accelerate the World's Progress on Nutrition. Washington, DC: 2014.

[3] International Food Policy Research Institute. *Global Nutrition Report 2015: Actions and Accountability to Advance Nutrition and Sustainable Development*. Washington, DC: 2015.

[4] International Food Policy Research Institute. Global Nutrition Report 2016: From Promise to Impact: Ending Malnutrition by 2030. Washington, DC: 2016.

Table 1: Overview of evolution for the monitoring rules to assess country progress towards the global nutrition targets

Indicator	WHO proposal and GNR 2014 (p.20)	GNR 2015 (p. 12-13) and GNR 2016 (p. 117)
Stunting	**On course** Current AARR[a]>= required AARR[b] **Off course** Current AARR[a]< required AARR[b]	**On course: Good progress** Current stunting rate $\leqslant$ 5% and current AARR[a] $\geqslant$ 0 (the stunting rate is 5% or below and declining further), or current AARR[a] $\geqslant$ country-specific AARR required to meet global goal, irrespective of prevalence (rate of decrease is faster than or equal to rate needed to meet global goal); **On course: At risk of off course** Current stunting rate $\leqslant$ 5% and current AARR < 0 (stunting rate is 5% or below, but increasing) **Off course: Some progress** Current stunting rate > 5% and current AARR > 0, but < country-specific AARR required to meet global goal (the stunting rate is above 5% and declining, but not fast enough to meet global target) **Off course: No progress** Current stunting rate > 5% and current AARR $\leqslant$ 0 (the stunting rate is above 5% and stationary or getting worse)
Anaemia	**On course** Current AARR[c]>= required AARR[b]; **Off course** Current AARR[c]< required AARR[b];	**On course** Current AARR[c] >= 5.2 **Off course** Current AARR[c] < 5.2
Low birth weight	Not available	Not available
Overweight	**On course** No increase in prevalence compared with country baseline and <7%[d] **Off course**	**On course: Good progress** Current overweight rate < 7% and current AARR[a] $\geqslant$ 0 (overweight rate is below the 7% threshold and decreasing); **On course: At risk** Current overweight rate < 7% and current AARR[a] < 0 (overweight rate is below the 7% threshold but increasing) **Off course: Some progress** Current overweight rate $\geqslant$ 7% and current AARR[a] > 0 (overweight rate is above the 7% threshold but

Indicator	WHO proposal and GNR 2014 (p.20)	GNR 2015 (p. 12-13) and GNR 2016 (p. 117)
	Increase in prevalence compared with country baseline or ≥7%[d]	decreasing); **Off course: No progress** Current overweight rate ≥ 7% and current AARR[a] ≤ 0 (overweight rate is at or above the 7% threshold and increasing)
Exclusive breastfeeding	Not available	**On course** Current AAPPI[e] ≥ target AAPPI[f] **Off course: Some progress** Current AAPPI between 25% and 100% of target AAPPI[f]I; **Off course: No progress** Current AAPPI (positive or negative) is < 25% target AAPPI[f], and there is no decrease in exclusive breastfeeding rates of 10 percentage points or more; **Off course: Reversal** A greater than 10-percentage-point decrease in exclusive breastfeeding rates has taken place over any recent time period at any exclusive breastfeeding level;
Wasting	**On course** Wasting < 5%; **Off course** Wasting >= 5%;	**On course** Current wasting prevalence < 5% **Off course** Wasting rate ≥ 5%

[a] Current average annual rate of reduction (AARR calculated based on a log-linear regression using either the historical available data between 1999 and 2012 whenever there are no data beyond the baseline, or using the latest available estimate and the baseline (two points only – exponential growth formula) when latest estimate year is post-baseline.

[b] Required AARR for the country refers to the required rate of progress to the equivalent to the global target definition;

[c] Current AARR for anaemia is presently calculated based on the model-based estimates using two points only (latest year, i.e., 2011, and year 2000) with the exponential growth formula.

[d] The global estimate for overweight has since then been updated, and the current level replacing the previous 7% is 6%;

[e] Current average annual percent point increase (AAPPI) is calculated as the difference in rates divided by the number of years in the considered period, either using historical available data between 2005 and 2012, whenever there are no data beyond the baseline, or using the latest available estimate and the baseline when latest estimate year is post-baseline.

[f] Target AAPPI: countries with baseline EBF rates close or above 50% (the global target) are encouraged to continue efforts to increase EBF rates by a minimum AAPPI of 1.2%. If the required calculated AAPPI to reach 50% rate by 2025 is higher than or equal to 1.2, then the calculated value is used as the suggested AAPPI; otherwise the suggested AAPPI is: 1.2 if baseline rate is less than 60%, by 0.06 if baseline rate is between 60% and 70%, or equal to zero if baseline rate is higher than or equal to 70%.

Key issues and TEAM recommendations

In February 2016, WHO and UNICEF presented TEAM with a background document describing outstanding method issues for monitoring the six targets. TEAM made the following recommendations:

- As much as possible, maintain a cohesive approach across indicators.
- Place countries in three categories if possible, but allow for flexibility when drafting the rules.
- Collapse the ratings into two categories ("on track" and "off track", collapsing all other categories) when summarizing for regional or global assessment.
- Investigate the size of the confidence intervals for trend estimates to define a category for "no change".
- Incorporate information on both progress made and current prevalence for all indicators except wasting.
- Use AARR to measure progress for all indicators, by switching EBF indicator to "% not exclusively breastfed" (instead of percentage point increase in exclusive breastfeeding).
- Use all data points in the data series defined for current progress assessment (rather than baseline to endpoint).
- Devise rules that imply stricter national than global targets where appropriate.
- Apply, where relevant, relative global and national targets, except for those countries that have achieved a "minimum" acceptable prevalence level, which would be considered "on track".
- Study the number of countries affected by any proposed cut-offs that define categories.

Based on TEAM suggestions before the meeting, a "straw man" proposal for classifying country progress was presented based on the principle of a harmonized approach to monitoring across all indicators, to the extent possible. It was based on four categories for all indicators except wasting (two categories) and required that the country-level progress rates were aligned with those required at global level for each indicator. This proposal generated ideas about how to refine the algorithm and apply the approach to the global databases. The pending issues concerned: (1) level of "acceptable prevalence", (2) required Average Annual Rate of Reduction (AARR) with respect to the global (or recommended) targets, (3) a buffer around zero for "no progress", (4) the time window for assessing current (or recent) progress, and (5) the number of categories.

The following section presents conceptual material, a summary analysis of results when available, discussion and TEAM's recommendations for each of the five issues described above requiring further discussion.

(1) Acceptable prevalence level

For each indicator, it was proposed that countries should be considered "on track" if they had already achieved a low prevalence of the nutritional deficiency, regardless of the direction or magnitude of recent trends. TEAM originally proposed that these levels be

linked to levels of public health significance for all indicators. However, due to a lack of data to define public health significance, the group decided, instead, to select cut-offs with other purposes.

It was proposed that a 5% threshold be used for all indicators except for exclusive breastfeeding and childhood overweight. For exclusive breastfeeding, it was determined that a 30% threshold of not exclusively breastfed would appropriately reflect a feasible level of achievement (see further explanation on alternative thresholds considered under issue 2 below).

Following discussion on proposed rules for assessing country progress on overweight and their implications, the need to align those with the global target, i.e. no increase in overweight, was stressed. It was then proposed that there should be no acceptable prevalence level for the on-track overweight category. Even countries with low prevalence of overweight at baseline should be concerned if this prevalence is rising.

(2) Required Average Annual Rate of Reduction (AARR)

If the country prevalence level is beyond the "no concern" threshold specified above, then the current AARR, based on recent trends, will be compared to the required AARR to reach a set target for that country. TEAM was asked to advise on what the targets should be for countries to have comparable assessment across all countries. The intent was to use, as much as possible, consistent approaches across the indicators. It is worth reminding that these targets will be used for global assessment of countries' progress, or their contributions to the global progress towards targets, and do not necessarily align with countries' own national targets.

The following indicator-specific values were proposed as the required progress rates, mirroring, as much as possible, the required progress rates at global level:

o   Stunting: Globally, the target is a 40% reduction in the number of stunted children between 2012 and 2025. Achievement is thus dependent on both, prevalence of stunting and population growth. Taking into consideration the population growth and the target reduction between baseline and target year, one can calculate the target prevalence for 2025. Following, the AARR on the stunting prevalence required to reach the target, prevalence can be derived. For example, if there were no population growth, the 40% reduction would translate into a required AARR of 3.9% per year, but if there is population growth, the required AARR will be higher than that. It was then proposed that, for the rules drafted in this report, the required AARR for the country to reach a 40% reduction between 2012 and 2025 as just described be applied to all countries with a prevalence above 5%.

o   Anaemia: The global target is a 50% reduction in prevalence of anaemia in women of reproductive age between 2012 and 2025. Unlike stunting, this reduction is independent of population growth and is a straightforward reduction in prevalence. The proposed average annual rate of reduction, or AARR, required for the country to reach a 50% reduction in prevalence of anaemia in women of reproductive age between 2012 and 2025, which is 5.2% per year, should be applied to all countries with anaemia prevalence above 5%.

o   Low birth weight: The global target, that is a 30% reduction in the prevalence of low birth weight, translates into a required AARR of 2.74%. The proposed required AARR for low birth weight is thus 2.74%, to be applied to all countries with prevalence above 5%.

o   Exclusive breastfeeding: The global target is at least a 50% prevalence of exclusive breastfeeding by 2025, a fixed threshold instead of a relative increase, making it different from those for stunting, anaemia and low birth weight. As mentioned earlier, TEAM recommended that prevalence of not exclusively breastfeeding (NEBF) be used for drawing the rules for assessing country progress towards the EBF global targets and thus NEBF will be used hereafter. Although the global target is not based on a relative decrease in NEBF, it was proposed that, to translate the global target to country targets, a relative reduction in NEBF would be appropriate. In this way, countries with low prevalence of EBF are expected to make the greatest gains in EBF. At the global level, to reduce NEBF prevalence from 62% (the global 2012 estimate) to 50%, a 20% relative reduction in NEBF (or an AARR of 1.65%) would be required. One proposal was to apply this required AARR across all countries when considering the country progress assessment rules. This rule was considered in combination with a 10% "acceptable prevalence level," implying that countries should achieve a 20% relative reduction in NEBF until they reached an EBF prevalence of 90%. However, TEAM considered that a 20% change in NEBF over a 13-year period is small compared to the other targets of 30%, 40%, and 50% change, especially given that there are known and effective interventions to increase breastfeeding rates, and conversely, that a 90% prevalence of exclusive breastfeeding is extremely high (no country has achieved this). Thus, TEAM proposed to simultaneously increase the required rate of reduction in NEBF from 20% to 30% by 2025 (translating to a required AARR of 2.74%) and set the acceptable prevalence level at 30%. This increases the level of ambition when applied to countries but focuses that ambition on countries with lower baseline prevalences.

o   Overweight: The target for overweight was to have no increase in the prevalence. The first proposal presented to TEAM departed from the global target. It was because some countries have high prevalence of overweight in children under 5

years of age, and thus a reduction in prevalence, of 40%, should be considered when drafting the rules for country progress assessment across all countries. However, TEAM raised concerns about using monitoring rules for classifying countries as on- or off-track that are not aligned with global targets, especially in this case, when at global level, zero decrease is required. Following these discussions, TEAM proposed that the on-track category be based on the "no increase" required and that this be applied to all countries.

o <u>Wasting</u>: The global target for wasting is to reduce and maintain the prevalence below 5%. The source for this indicator is surveys, which are carried out in different periods and, sometimes, under diverse circumstances. Child wasting can change rapidly, affecting resulting rates and making it difficult for estimates to be comparable over time. Thus, TEAM recommended exploring ways of assessing countries' progress other than regular trends. One approach tested was that of moving averages of wasting, to smooth out seasonal variations. The analyses, however, demonstrated that due to commonly large time periods between surveys (at least 3 to 5 years apart), smoothing was not effective for interpreting progress (results not shown). Ideally, smoothing would help if in the same year there were multiple surveys, covering potential seasonal effects that cause fluctuations. Nevertheless, there is value in recognizing countries that are making progress in consistently decreasing their wasting prevalence. TEAM recommended that an intermediate category be created to identify countries that show a tendency toward decreasing the prevalence of wasting over multiple years (see next section for more details).

(3) Establishing a buffer around zero for "no progress"

Although the cut-off zero provides a straightforward division between "above zero" (making progress) and "below zero" (moving in the wrong direction), there are strong arguments for avoiding it. First, given the inaccuracy in estimating the rate of progress based a small number of surveys which have their own intrinsic variability, many countries would be misclassified in trying to separate those making some progress versus those stagnating and moving in the wrong direction. Conceptually, it might make sense to group countries with an AARR of "around zero" because these are all "stagnating" countries with little or no progress. Second, there is precedent for the use of a cut-off different than zero to distinguishing those countries clearly showing progress from those that do not (e.g., Millennium Development Goal monitoring for underweight as well as other sectors such as Water and Sanitation and mortality used a cut-off different from zero).

In their meeting in February 2016, TEAM agreed that a "typical" confidence interval around zero should be defined for each indicator to delineate the category of "no progress", for all countries. Given the scarcity of data for each country and for most indicators, however, confidence intervals using standard errors based on country's data would lead to very wide, meaningless ranges, overlapping greatly with the other progress assessment categories.

Using bootstrap sampling, standard errors (SEs) were calculated for the AARR of stunting, overweight, and exclusive breastfeeding (see details in Annex 1). For trends starting in 2005, confidence intervals would be very wide. For example, for an indicator as stable as stunting, the estimated standard error of 1.46 suggests that the confidence interval around zero to define the category of "no progress" would be approximately ± 2.9% in the best-case scenario. Given that "on track" would use a cut-off of roughly 3.9%, this would make the "some progress" category very limited. For overweight, the 95% CI implied by the calculated SE would place even large drops in overweight into the "no progress" category. The same would be true for exclusive breastfeeding. Thus, the high SEs indicate that a 95% CI would not be useful for monitoring country progress.

Comparing the SE results across indicators suggests that the uncertainty of the AARR is much greater for overweight than for stunting, typically about three times larger. The SEs are somewhat larger for exclusive breastfeeding than for stunting. Thus, it may be prudent to use proportionally wider buffers for overweight and exclusive breastfeeding in defining the "no progress" category.

Given that a truly 95% confidence interval approach is not practical, a more pragmatic approach for helping countries could be a descriptive one – as was done for the MDGs, when monitoring child underweight, just mentioned above. While admittedly arbitrary, the limits of the "no progress" category were proposed at ± 0.5% for stunting, LBW and anaemia. For overweight, the typical interval ± 1.5% was proposed, based on the ratio of SEs for overweight vs. stunting shown in Annex 1. Similarly, for exclusive breastfeeding, the typical interval ± 0.8% was proposed. While there is no empirical result on the SEs of the AARR for wasting, it is anticipated that its variability would be even larger due to seasonal fluctuations and the volatility in emergency settings. The proposed typical interval for wasting was ± 2.0%, resulting in a wide range for the category "no progress". For global monitoring only, it was proposed that the classification "off track" of the AARR for wasting be divided into three in recognition of the steady progress of some countries in reducing acute malnutrition. TEAM agreed that these typical intervals, as proposed, would provide buffers around zero.

(4) Time window for assessing progress

In the GNR, recent trends for monitoring the nutrition targets were estimated based on two approaches depending on whether a country had collected data after its baseline year. For those countries with post-baseline data, progress was evaluated from the baseline year onwards. For those with no post-baseline data, historical data (before the baseline) were used. At its February 2016 meeting, TEAM considered that this dual approach would produce inconsistencies in assessments between countries, and it thus proposed alternatives.

One option was to report only on progress for those countries with post-baseline data. This

would leave a number of countries classified as having "no information," although the number would be expected to decline over time. A second option would be to shift the baseline to sometime before 2012 for all countries and show trends from an earlier point as had been done for the MDGs. Progress could be assessed across all countries by using data starting from a year before 2012.

In September 2016, TEAM reviewed four options for defining the underlying data series for calculating the current/recent AARR:

1. Use the whole series of data available from the starting year 2005. Progress would be assessed for all countries with a prevalence level lower than the "acceptable level," or with at least two data points since 2005, regardless of whether there are any post-2012 data.
2. Use the whole series of data available from the starting year 2005 as in approach 1 but assess progress only for countries with at least one post-baseline data point.
3. Use the whole series of data available from the starting year 2008. Progress would be assessed for all countries with prevalence levels lower than the "acceptable level," or with at least two data points since 2008, regardless of whether there are any post-2012 data.
4. Use the whole series of data available from the starting year 2008 as in approach 3 but assess progress only for countries with at least one post-baseline data point.

TEAM agreed that 2008 should be used as the starting year for assessing recent trends and stressed the importance for countries to have post-baseline data for measuring recent progress. The group recommended that, for global monitoring purposes, no progress assessment classification should be done for countries where data are unavailable beyond the baseline. They also highlighted the importance of collecting nationally representative data every three years to provide reliable trend estimates.

(5) Number of categories

There were two options for the number of categories and labels presented to TEAM:

Four categories: 1. On-track 2. Some progress 3. No progress 4. Worsening

Three categories: 1. On-track 2. Some progress 3. No progress or worsening

Weighing pros and cons, TEAM recommended using three categories for all indicators (plus a "No data available" designation), consistent with MDG monitoring for underweight. For overweight, TEAM agreed to have only two categories: (1) "on track" when current AARR is equal to or above the 1.5 buffer (this category includes those countries clearly showing progress and those within the "stagnation" interval) and (2) "off track" (for those countries clearly facing a worsening situation, or increased prevalence).

Monitoring methodology recommended by TEAM for assessing country progress

A summary the country progress monitoring assessment rules as recommended by TEAM is presented in Table 2.

Table 2. Proposed monitoring rules and classification of progress towards achieving the six nutrition targets

Indicator	On track	Off track - some progress	Off track - no progress or worsening
Stunting	AARR $\geq$ required AARR [a] or level < 5%	AARR < required AARR [a] but $\geq$ 0.5	AARR < required AARR [a] and < 0.5
Anaemia	AARR $\geq$ 5.2 [b] or level < 5%	AARR < 5.2 but $\geq$ 0.5	AARR < 0.5
LBW	AARR $\geq$ 2.74 [c] or level < 5%	AARR < 2.74 but $\geq$ 0.5	AARR < 0.5
NEBF [d]	AARR $\geq$ 2.74 [e] or level < 30%	AARR < 2.74 but $\geq$ 0.8	AARR < 0.8
Wasting	Level < 5%	Level $\geq$ 5% but AARR $\geq$ 2.0	Level $\geq$ 5% and AARR <2.0
	On track	**Off track**	
Overweight	AARR $\geq$ -1.5	AARR < -1.5	

[a] Required AARR based on the stunting prevalence change corresponding to a 40% reduction in number of stunted children between 2012 and 2025, considering the estimated population growth estimated (based on data from UN Population Prospects).
[b] Required AARR based on a 50% reduction in prevalence of anaemia in women of reproductive age between 2012 and 2025.
[c] Required AARR based on a 30% reduction in prevalence of low birth weight between 2012 and 2025.
[d] NEBF stands for the prevalence of not exclusively breastfed.
[e] Required AARR based on a 30% reduction in not exclusively breastfed rate between 2012 and 2025.

Using the latest available datasets for each of the indicators, countries with sufficient data were classified in keeping with the categories presented in Table 2. For low birth weight, work is being carried out by an inter-agency/institution group of experts and national estimates. The distribution of the number of countries in each category for each indicator is presented in Table 3. The number of countries for which a determination can be made is relatively low (fewer than 50) for all the indicators, except for anaemia, for which model-based estimates are available for every year between 1990 and 2016.[5] This results from the decision to require at least two data points since 2008, one of which must be since 2012 (see issue 4 above). The number of countries that can be classified will increase over time as more

---

[5] Anaemia estimates for countries in the WHO European region were undergoing validation, and thus subject to changes, at the time this report was written.

data are collected. For all indicators except anaemia, about 40-50% of the countries for which assessment was provided fall in the "on track" category and there is a distribution of countries across "some progress" and "no progress or worsening". This implies that the proposed rules will effectively differentiate among countries with diverse experiences in achieving targets. For anaemia, no country was assessed as "on track", and only thirty percent of those off track presented some progress. An important note is, however, that despite a high proportion of countries having nationally representative survey data available for anaemia, there is still a lack of reporting on this indicator, especially in high-income countries. As a result, estimates may not capture the full variation across countries and regions, tending to shrink towards global means when data are sparse. Moreover, although model estimates are provided for countries for every year between 2000 and 2016, survey data on haemoglobin levels were available for 116 countries, out of which only 30 countries had data beyond the baseline (2012). Nevertheless, country trends confirm little or no progress for the great majority. This reinforces the fact that further investigation on the determinants of anaemia in women of reproductive age is required, followed by the adoption by countries of recommendations for effective implementation of nutrition sensitive and specific interventions. In addition, the monitoring of the implementation of interventions will be essential to ensure their efficacy.

Table 3. Classification of countries using the proposed monitoring rules

Indicator	# countries classified	On track	Off track - some progress	Off track - no progress or worsening
Stunting [a]	44	17	19	8
Anaemia [b]	189[c]	0	49	140
NEBF [d, e]	45	21	8	16
Wasting [a]	49	25	7	17
		On track	Off track	
Overweight [a]	31	13	18	

[a] Data source: UNICEF-WHO-WB Joint Child Malnutrition Estimates 2017 edition dataset; [b] WHO Global Health Observatory. Geneva: World Health Organization; 2017; [c] Only , 30 countries have at least one survey data point later than 2012; results of this classification should be interpreted with caution. [d] Data source: UNICEF Database on Infant and Young Child Feeding - 2016 edition dataset on exclusive breastfeeding; [e] NEBF stands for the prevalence of not exclusively breastfed.

Conclusion

This report provides important guidance on how to classify countries based on progress in achieving the nutrition targets using a comparable global approach. The methodology described here are intended for use in global monitoring reports to WHO governing bodies, in WHO/UNICEF global publications, and in the GNR. It is thus envisaged that it will be published and widely distributed, especially among partners involved in global monitoring efforts. It is probable that it will require refinement as more data become available and countries provide feedback.

The guidance provided is based on comparable country targets for tracking global progress in the light of national levels and trends. However, they are not intended to replace countries' own targets and monitoring practices, although countries could adapt them for national monitoring based on their own targets.

The proposed guidelines for tracking country progress toward the nutrition targets effectively implement key principles the TEAM considered to be important. First, they incorporate information on both the current level of each indicator and recent trends. Second, they use a common metric, the average annual rate of reduction, for all indicators to describe whether continued trends in the future will lead to target attainment. Third, they hold all countries to a comparable standard based on a single global target by applying the same relative reduction across all countries. Fourth, they acknowledge countries that are making progress toward the target, even if the rate is insufficient to forecast that the target will be achieved.

The deliberations on a tracking algorithm also raised clear concerns about the global architecture of nutrition surveillance. Too few countries regularly collect nationally representative data to provide reliable trend estimates. Ideally, indicators should be available for all countries at least every three years. Furthermore, adherence to global data collection standards is critical to supplying comparable estimates across countries.

The Sustainable Development Goals have underlined the importance of ending all forms of malnutrition by 2030. They reiterated the targets for nutrition while highlighting the importance of holding countries accountable for progress towards their achievement. The work described in this report represents the thinking of technical experts on tracking progress on a fair, feasible, comparable, and scientifically sound basis. It is hoped that the guidelines will serve as a useful tool for eliminating malnutrition.

Annex. Estimation of standard errors of the current AARR estimates

The purpose of this exercise was to estimate the degree of uncertainty around the estimates of the current Average Annual Rate of Reduction (AARR) for stunting, overweight and exclusive breastfeeding. For stunting and overweight, the 2015 edition of the joint malnutrition estimates (JME)[6] dataset was used (total of 149 countries), while for the exclusive breastfeeding, the dataset shared by UNICEF[7] in July 2015 was used (total of 139 countries).

The current AARR is calculated using a log-linear regression (exponential growth model) following the approach used by UNICEF for monitoring Millennium Development (MDG) related to the indicator underweight.[8] In sum, the formula for its calculation is given by *AARR=1 - exp(β)*, where β is the slope of the natural logarithm of the prevalence regressed on the survey year. For the present analyses, four potential starting years were considered to defined data series: 1985, 2000, 2005 and 2008. When only two points are available, the formula simplifies to *AARR = $ln(P_t/P_{t+n})/n$*, where P_{t+n} is the final prevalence in the data series, P_t is the starting prevalence, and *n* is the number of years between the two data points.

Note: although the stunting target focusses on number of stunted children (40% reduction between 2012 and 2025), its monitoring is based on the change in prevalence plus considering the population growth (according to the UN Population Prospects latest estimates).

Three approaches for estimating the standard error (SE) around the current AARR estimate were considered for stunting (note we chose this indicator given that it one with most available data):

- Approach 1: the AARR SE was calculated based on the log-linear regression as described above, using the SE of the slope (*β*) followed by a back transformation to the AARR scale (Delta method)[9].
  That is, if *AARR = 1 - exp(β)* then *SE(AARR) = exp(β) SE (β).*
  The standard error of the slope (simple linear regression methodology) is only available when country to have 3 or more data points. A Gamma distribution was then fitted to the set of available SE(AARR) estimates, after removing clear outliers, based on what analysis was carried out performing bootstrap sampling to estimate the median SE of

---

[6] UNICEF-WHO-WB Joint Child Malnutrition Estimates 2015 edition dataset.

[7] UNICEF Database on Infant and Young Child Feeding - 2015 edition dataset on exclusive breastfeeding.

[8] UNICEF Division of Policy and Practice, Statistics and Monitoring Section (2007) Technical Note: How to calculate Average Annual Rate of Reduction (AARR) of Underweight Prevalence. UNICEF: New York.
http://www.childinfo.org/files/Technical_Note_AARR.pdf (Accessed 29 July 2013).

[9] Casella, G.; Berger, R. L. Statistical Inference (2nd ed.). Duxbury Press 2002. 688p.

the AARR and the confidence interval around it (2.5th and 97.5th quantiles). For each considered dataset, 10,000 samples of the same size as the original dataset after removing the outliers were generated for deriving median and confidence interval estimates.

- Approach 2: the entire data series in the JME dataset for all countries was used (year 1985 onwards), and 10-year snapshots were extracted to calculate series of AARRs for each country. Based on those, standard deviations (SD) were then calculated for each country and the distribution of these SDs was explored. The 10-year series was selected to mimic more closely what one would expect to be a sufficient time range to monitor countries progress. Like approach 1, a Gamma distribution was fitted and bootstrap sampling were draw to estimate the median, and the 2.5th and 97.5th quantiles. Figure 1 illustrates the different AARR's for each country, depending on the 10-year-serie snapshot considered for the estimation. Bootstrap sample from the fitted Gamma distribution (10,000 samples of size 105) provided the median AARR SD estimated at 3.2 (95% CI 3.7; 4.3).

- Approach 3: Based on a set of surveys that were re-analysed to include SEs of the prevalence estimates (233 surveys for 93 countries), the AARR SE estimates were calculated based on only two points (from 2000, the earliest and the latest surveys). Based on the 2-points-based simplified formula above, that is, AARR=$\ln(P_1/P_2)/(t_2-t_1)$, the Delta method was used to approximate the SE of the transformed prevalence as $SE(AARR) = \sqrt{SE^2(\ln(P_1)) + SE^2(\ln(P_2))}/(t_2 - t_1)$, where P_1 and P_2 are the starting and ending prevalence estimates, respectively, and t_1 and t_2 the corresponding years. Like approach 1, a Gamma distribution was fitted and bootstrap sampling were draw to estimate the median SE, and the 2.5th and 97.5th quantiles. Out of 93 countries, SE estimates were derived for 59 countries (countries with two points). Bootstrap samples from the fitted Gamma distributions (10,000 samples of size 59) provided the median AARR SE estimated at 0.94 (95% CI 0.70; 1.21) based on this approach.

The first and third approaches were used to assess the variability around the mean AARR, while the second approach relates to the variability around the countries' AARR's.

For exclusive breastfeeding (EBF), the global target is to have the EBF rate increased to at least 50%. The linear point percent increase, that is, the Average Annual Point Percent Increase (AAPPI) is an option when monitoring country progress towards their own target. This measure is estimated as the slope of the linear regression line fitted to the data series. However, for the sake of consistency across the target indicators, the TEAM recommended that the use of the AARR of the non-exclusive breastfeeding be explored.

Hereafter, results from the first approach for different starting years will be presented for stunting, overweight and exclusive breastfeeding (or non-exclusive breastfeeding).
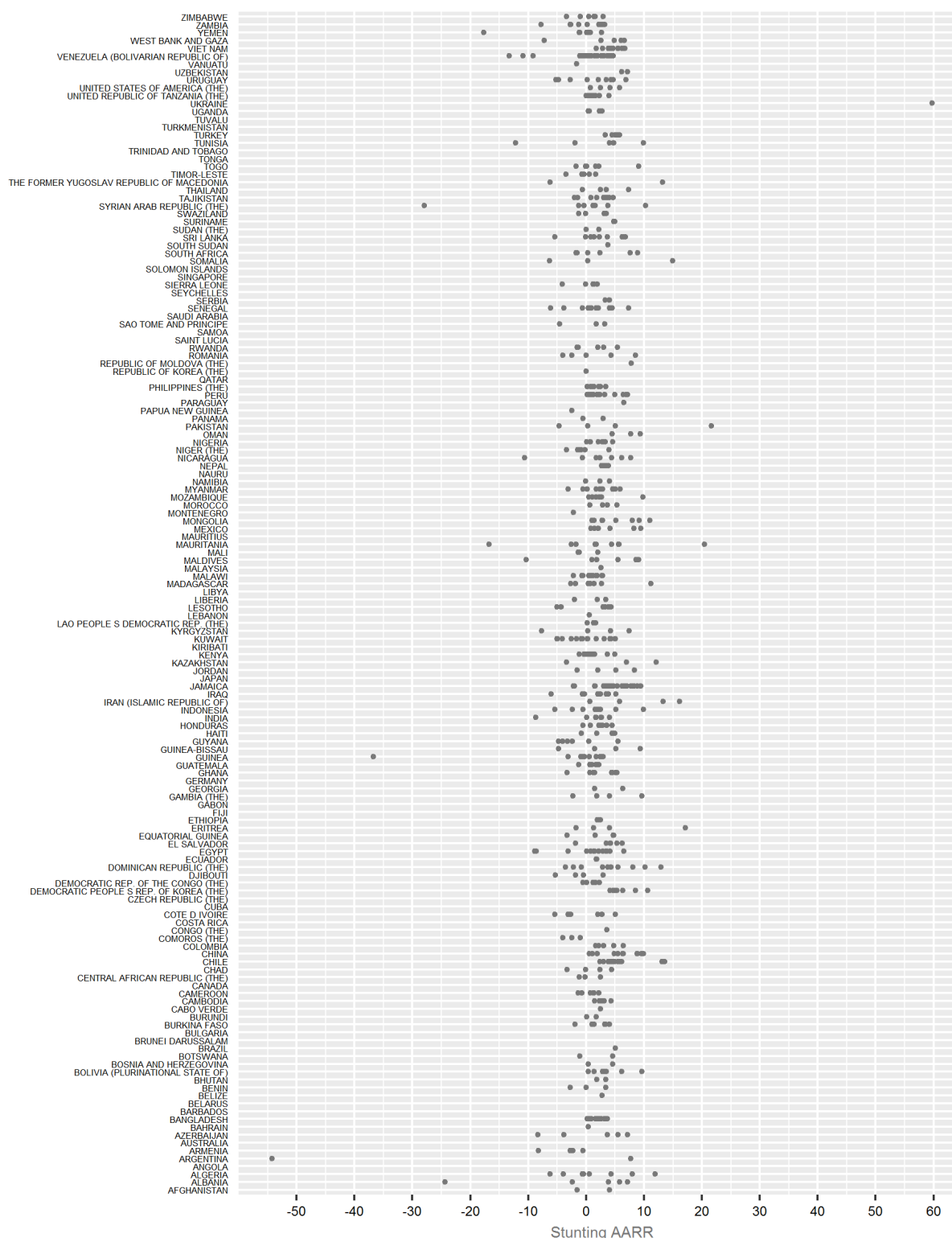
Figure 1. 10-year-snapshot Average Annual Rate of Reduction (AARR) estimates for stunting for each country.

Stunting

For each of the datasets with country data series starting at years 1985, 2000, 2005 and 2008, the AARRs and corresponding standard errors were calculated. Figures 2 (a) to (d) show the distributions of the AARR SEs for each of the starting years (1985 to 2008, respectively), and a Gamma distribution density (empirical Kernel fitted density - black line - and the Gamma density - red line).
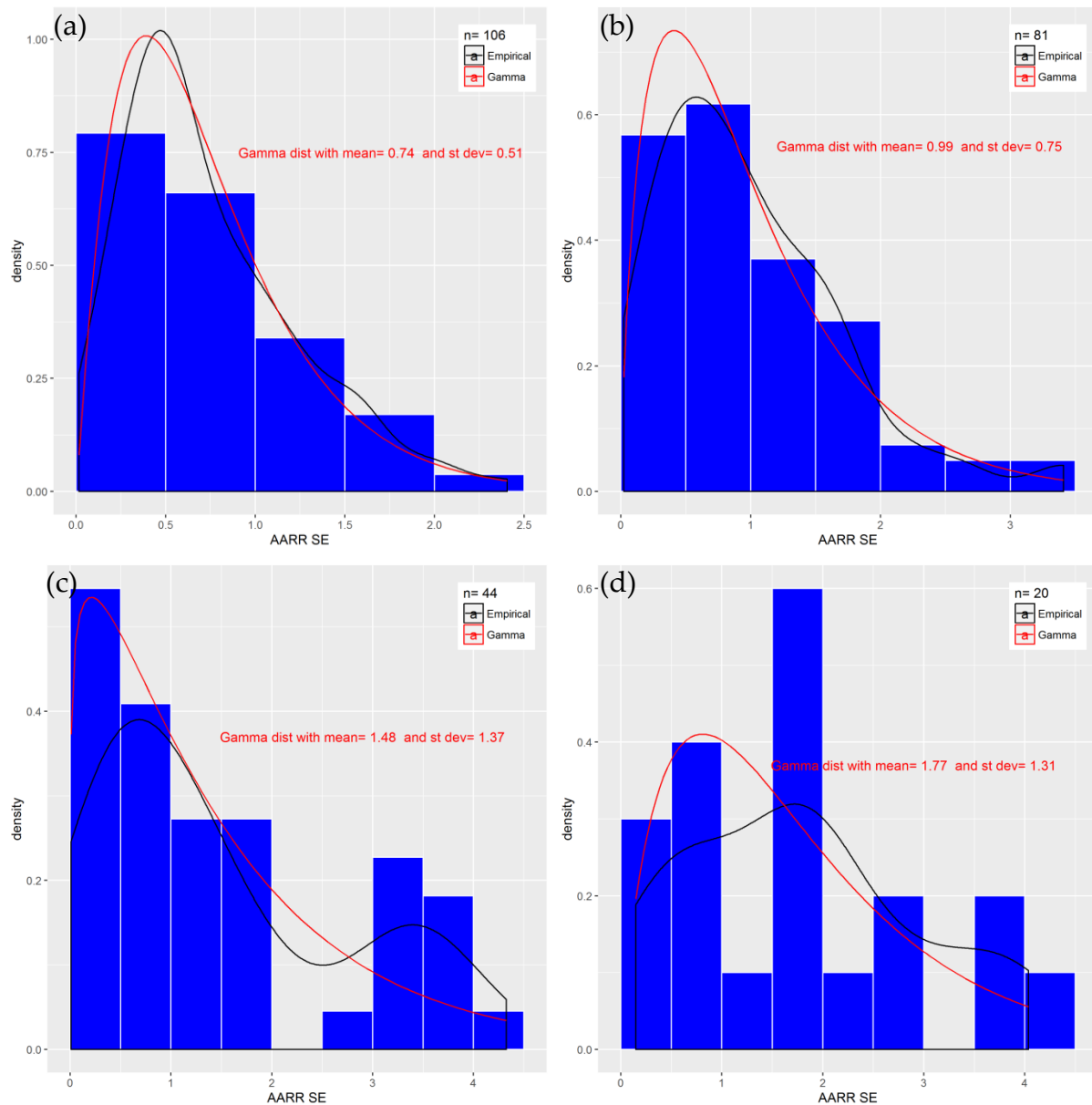

Figure 2. Distributions of the AARR SEs for stunting fitting a non-parametric Kernel density (labelled Empirical, black line), and a Gamma distribution density (labelled Gamma, red line), based on all data for stunting from 1985 (a), 2000 (b), 2005 (c) and 2008 (d).

Using the fitted Gamma distributions in the four data sets above, after removing the outliers, analysis was carried out based on bootstrap sampling to estimate the median SE of the AARR and the confidence interval around it (2.5th and 97.5th quantiles). For each case, 10,000 samples of the same size as the original dataset after removing the outliers, when existing (that is, 106, 81, 44, and 20) were withdrawn from the AARR SE fitted distributions and the 50th, 2.5th and 97.5th quintiles were calculated. Results are shown in Table A1.

Table A1. Median Average Annual Rate of Reduction (AARR) standard errors (SE) and the 2.5th and 97.5th quantiles for stunting based on their distributions, calculated from the country-specific log-linear regression models (bootstrap, 10,000 draws)

Starting year	total # countries	# countries with n⩾2	# countries with n⩾3	Bootstrapping Median AARR SE	95% CI	
					Lower	Upper
1985	149	134	110	0.74	0.64	0.84
2000	144	115	84	0.98	0.83	1.16
2005	133	93	44	1.46	1.10	1.90
2008	117	52	22	1.75	1.25	2.38

Data source: UNICEF-WHO-WB Joint Child Malnutrition Estimates dataset – 2015 Edition.

Overweight

The same four datasets as used for stunting, that is, years starting from 1985, 2000, 2005 and 2008 were considered to assess the AARR SE distribution for overweight. Figures 3 (a) to (d) show the distributions of the AARR SEs for all countries, fitting a non-parametric Kernel density, following data closely (labelled Empirical, black line), and a Gamma distribution density (labelled Gamma, red line) for the four datasets with different starting survey year.

Like what was done for stunting, based on the fitted Gamma distributions for each of the datasets (except for that starting in 2008), bootstrap estimates of the median SEs and 95% confidence intervals were derived. For each dataset, 10,000 samples of the same size as the original dataset (that is, 100, 76, and 39) were withdrawn from the AARR SE fitted distributions and the 50th (median), 2.5th and 97.5th quintiles were calculated. For the dataset with starting survey year 2008, sample size was too small and the gamma distribution not considered and only average AARR SE was calculated. Table A2 shows the results.
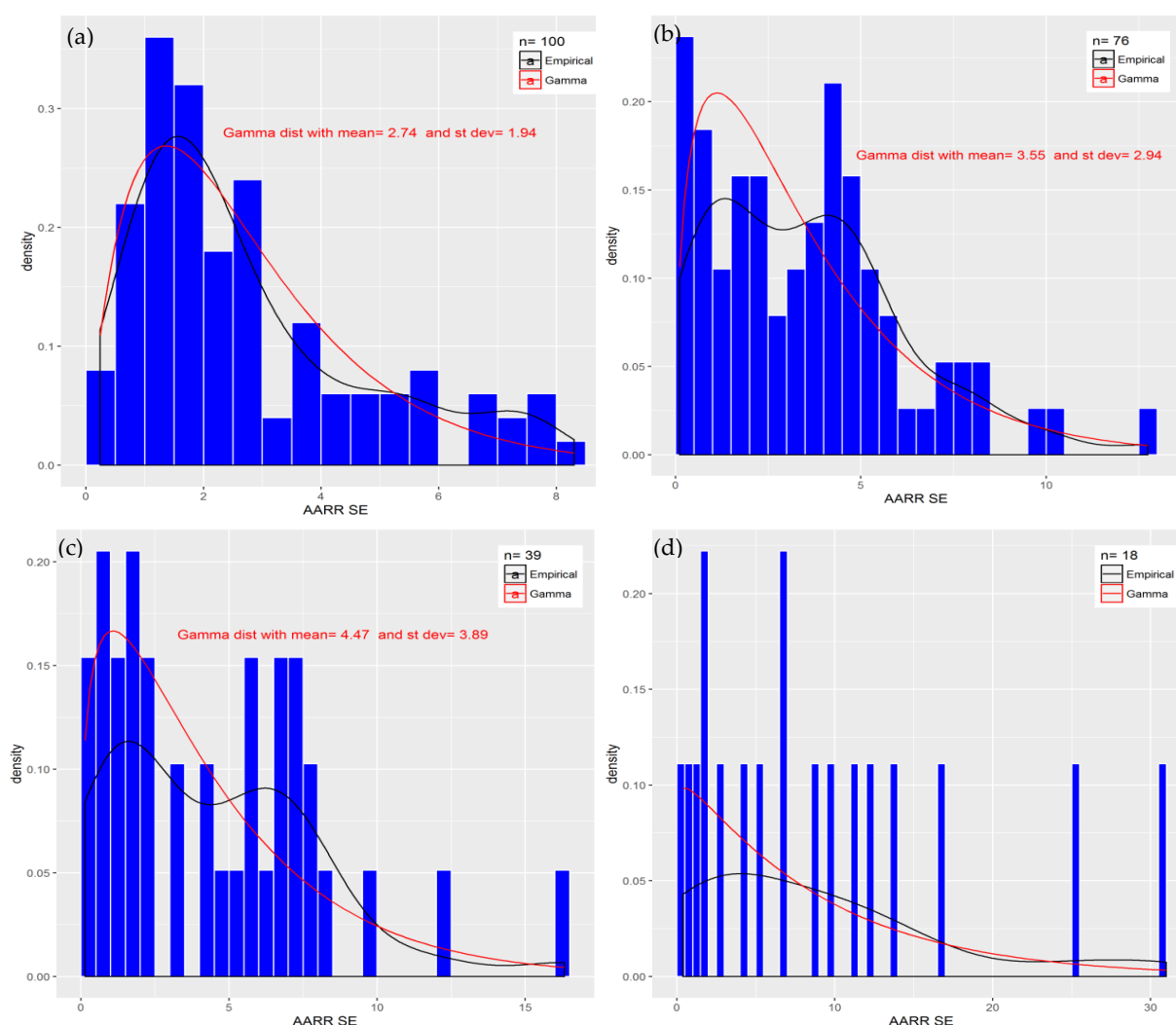
Figure 3. Distributions of the AARR SEs for overweight fitting a non-parametric Kernel density (labelled Empirical, black line), and a Gamma distribution density (labelled Gamma, red line), based on all data for stunting from 1985 (a), 2000 (b), 2005 (c) and 2008 (d).

Table A2. Median Average Annual Rate of Reduction (AARR) standard errors (SE) and the 2.5th and 97.5th quantiles for overweight based on their distributions, calculated from the country-specific log-linear regression models (bootstrap, 10,000 draws)

Starting year	total # countries	# countries with $n \geqslant 2$	# countries with $n \geqslant 3$	Bootstrapping Median AARR SE	95% CI	
					Lower	Upper
1985	149	127	100	2.73	2.37	3.13
2000	144	109	76	2.92	3.54	4.24
2005	133	87	39	4.43	3.33	5.77
2008	117	46	18	8.90*	NA	NA

Data source: UNICEF-WHO-WB Joint Child Malnutrition Estimates dataset – 2015 Edition; *Average AARR SE, not based on bootstrap but on the original sample of size 18.

Exclusive breastfeeding (EBF) and not-exclusively breastfed (NEBF)

Based on the latest dataset for the exclusive breastfeeding[10] at the time of these analyses (total of 139 countries), the same approach 1 used for stunting, based on the distributions of the AAPPI/AARR standard errors (SE) calculated from the country-specific log-linear regression models as described above, was selected.

Like stunting and overweight, four datasets considering years starting from 1985, 2000, 2005 and 2008 were considered to assess the AARR SE distribution for the two indicators. Figures 4 (a) to (d) and figures 5 (a) to (d) show the distributions of the EBF AAPPI SEs and NEBF AARR SEs, respectively, for starting years 1985, 2000, 2005 and 2008, respectively. A non-parametric Kernel density, following data closely (labelled Empirical, black line), and a Gamma distribution density (labelled Gamma, red line) were fitted for the four datasets with different starting survey year.
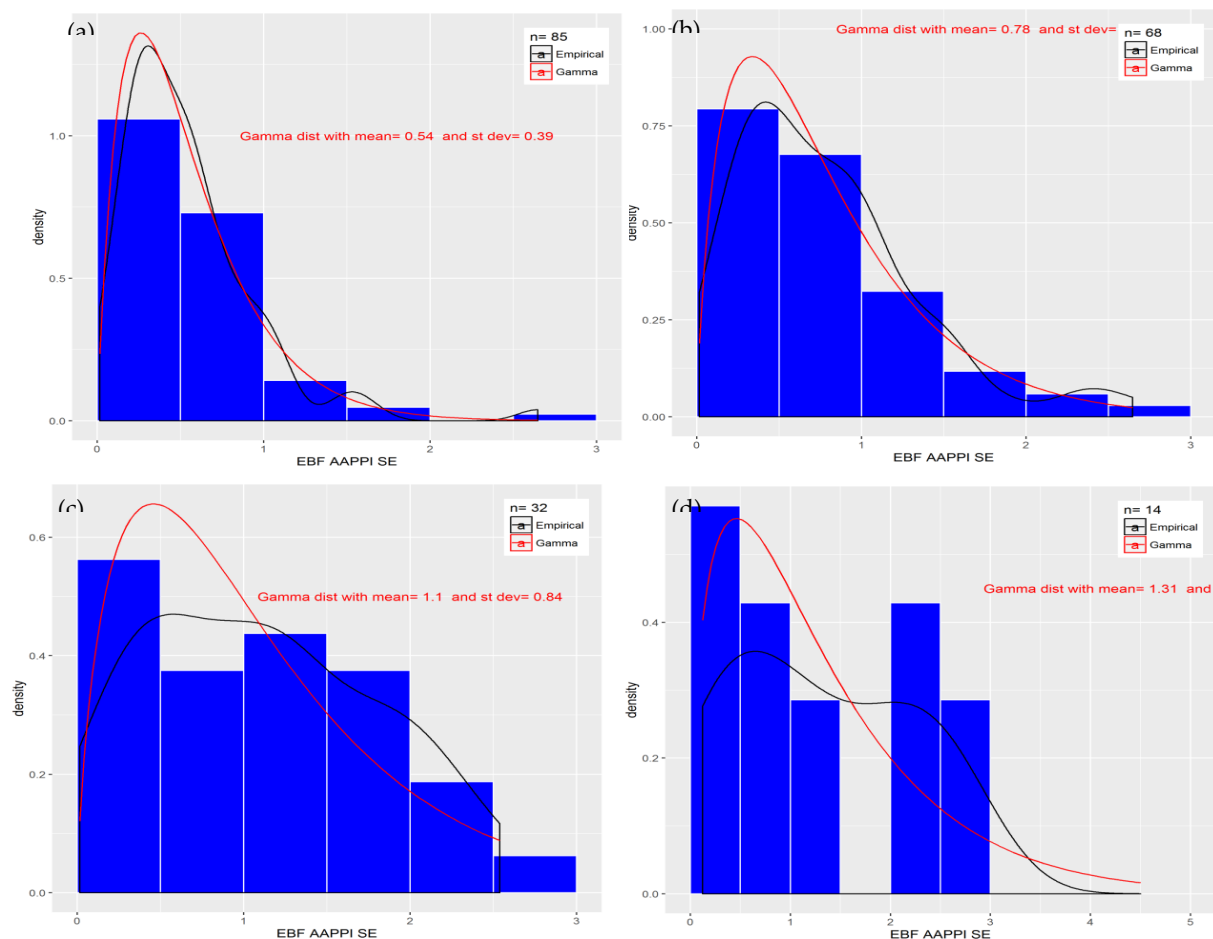

Figure 4. Distributions of the AAPPI SEs for exclusive breastfeeding fitting a non-parametric Kernel density (labelled Empirical, black line), and a Gamma distribution density (labelled Gamma, red line), based on all data for stunting from 1985 (a), 2000 (b), 2005 (c) and 2008 (d).

---

[10] UNICEF Database on Infant and Young Child Feeding - 2015 edition dataset on exclusive breastfeeding;
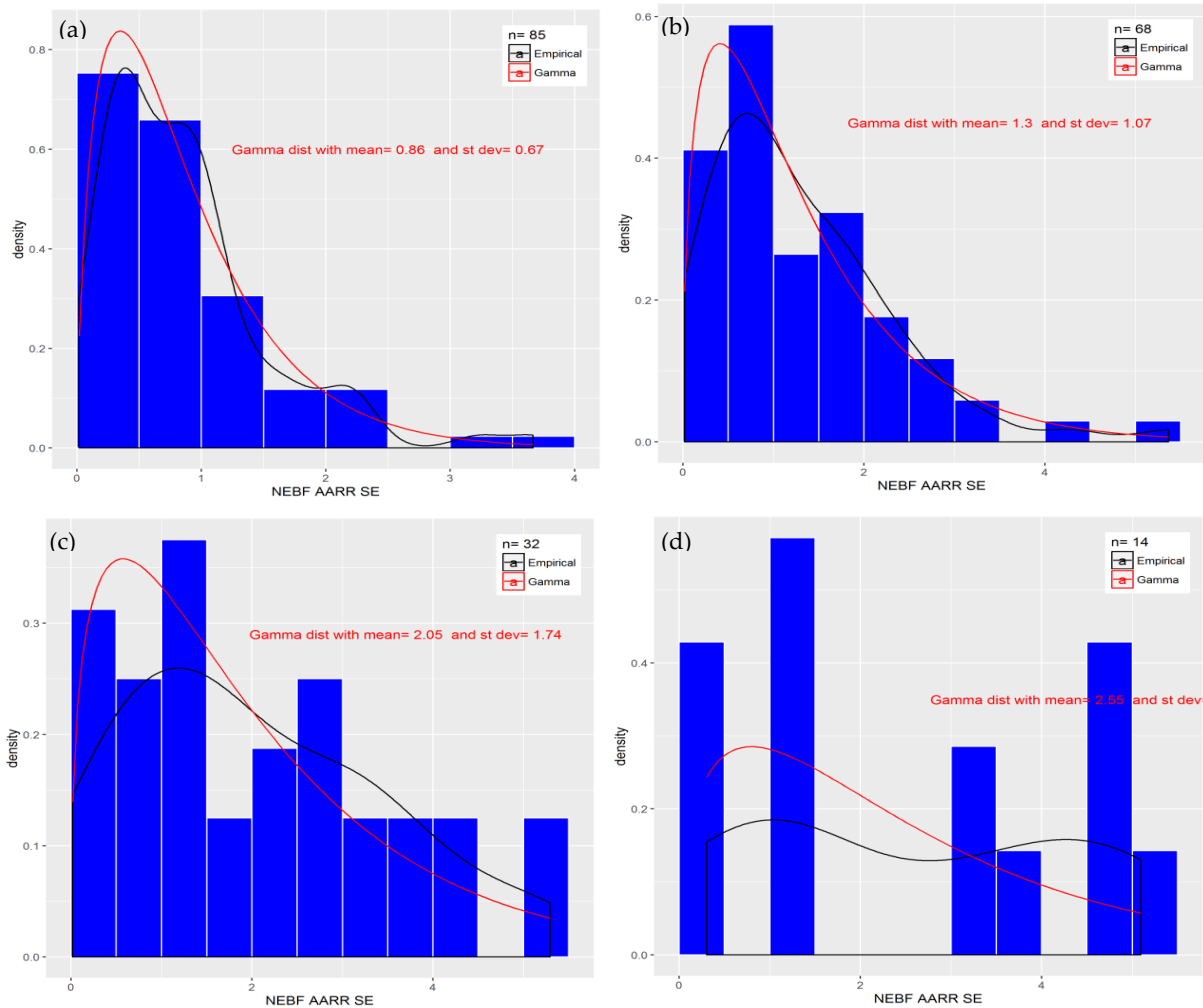
Figure 5. Distributions of the AARR SEs for non-exclusive breastfeeding fitting a non-parametric Kernel density (labelled Empirical, black line), and a Gamma distribution density (labelled Gamma, red line), based on all data for stunting from 1985 (a), 2000 (b), 2005 (c) and 2008 (d).

Like what was done for stunting and overweight, based on the fitted Gamma distributions for each of the indicators and datasets (except for that starting in 2008), bootstrap estimates of the median SEs and 95% confidence intervals were derived. For each indicator and dataset, 10,000 samples of the same size as the original dataset (that is, 100, 76, and 39) were withdrawn from the AAPPI SE and AARR SE fitted distributions and the 50th, 2.5^{th} and 97.5^{th} quintiles were calculated. For the datasets with starting survey year 2008, sample size was too small and the gamma distribution not considered and only average AAPPI SE and AARR SE were calculated. Table A3 shows the results.

Table A3. Median standard errors (SE) of the Average Annual Point Percentage Increase (AAPPI) for exclusive breastfeeding and of the AARR for no- exclusive breastfeeding, and the 2.5th and 97.5th quantiles based on their distributions, calculated from the country-specific log-linear regression models (bootstrap, 10,000 draws)

Starting year	total # countries	# countries with n≥2	# countries with n≥3	Bootstrapping Median		95% CI Lower	Upper
1985	139	110	85	AAPPI	0.54	0.46	0.62
				AARR	0.86	0.73	1.01
2000	131	101	68	AAPPI	0.78	0.65	0.93
				AARR	1.29	1.05	1.56
2005	124	82	33	AAPPI	1.10	0.83	1.42
				AARR	2.03	1.49	2.70
2008	109	38	14	AAPPI	1.31*	NA	NA
				AARR	1.83*	NA	NA

Data source: UNICEF Database on Infant and Young Child Feeding - 2015 edition dataset on exclusive breastfeeding; *Average AARR SE, not based on bootstrap but on the original sample of size 14.

Results presented in tables A1 to A3 confirm that the uncertainty around the AARR decreases as the data series extends over time. Although 30-year trends in data are not effective for evaluating the impact of nutrition policies and programmes put in place because of increasing commitment to nutrition, they can be considered asymptotic (that is, as more data become available).